



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2019

---

## **L'intégration du vocabulaire des Documents linguistiques galloromans dans le DEAF électronique**

Glessgen, Martin-Dietrich ; Dallas, Marguerite

**Abstract:** The exploitation of databases for lexical purposes has undergone extensive development during the first two decades of the 21st century. The methodological implications of this fruitful alliance between corpora and electronic dictionaries nonetheless deserve particular attention, the digital aspects as well as those pertaining to lexicology and metalexicography. The integration of the non-literary vocabulary of the Documents linguistiques galloromans into the database of the Dictionnaire Étymologique de l'Ancien Français (DEAFpré), currently in progress, provides an opportunity to illustrate the demands of a rigorous conception of computer-based lexicography (i. e. based on sound philological principles, ideally innovative, manual treatment of lexemes with a computer-assisted element, concentration on etymological, semantic and syntagmatic aspects). The case-study presented in this article focuses on the potential for interpretation afforded by corpus-based lexicography, its other main aim being an evaluation of the results and the cost in terms of time. The computerised system elaborated for this purpose can be extended to other projects, particularly those focusing on medieval Galloromania.

DOI: <https://doi.org/10.1515/lex-2019-0009>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-183993>

Journal Article

Published Version

The following work is licensed under a Publisher License.

Originally published at:

Glessgen, Martin-Dietrich; Dallas, Marguerite (2019). L'intégration du vocabulaire des Documents linguistiques galloromans dans le DEAF électronique. *Lexicographica*, 35:235-267.

DOI: <https://doi.org/10.1515/lex-2019-0009>

Martin Glessgen / Marguerite Dallas

# L'intégration du vocabulaire des *Documents linguistiques galloromans* dans le DEAF électronique

- |     |   |     |   |
|-----|---|-----|---|
| 1   | Idée de départ : un dictionnaire électronique basé sur des textes documentaires | 3.4 | L'intégration sémantique  |
| 2   | Les étapes préparatoires : édition et programmation                             | 3.5 | La rédaction d'articles lexicologiques dans le cadre du DEAFpré |
| 3   | Réalisation lexicographique   | 4   | Les DocLing et le Godefroy                                      |
| 3.1 | La lemmatisation  | 5   | Conclusions et perspectives                                     |
| 3.2 | L'importation des lemmes dans la base de données du DEAFpré                     | 6   | Références bibliographiques                                     |
|     |   | 6.1 | Monographies et articles  |
|     |   | 6.2 | Dictionnaires et bases de données                               |

**Abstract:** The exploitation of databases for lexical purposes has undergone extensive development during the first two decades of the 21st century. The methodological implications of this fruitful alliance between corpora and electronic dictionaries nonetheless deserve particular attention, the digital aspects as well as those pertaining to lexicology and metalexicography. The integration of the non-literary vocabulary of the *Documents linguistiques galloromans* into the database of the *Dictionnaire Étymologique de l'Ancien Français* (DEAFpré), currently in progress, provides an opportunity to illustrate the demands of a rigorous conception of computer-based lexicography (i. e. based on sound philological principles, ideally innovative, manual treatment of lexemes with a computer-assisted element, concentration on etymological, semantic and syntagmatic aspects). The case-study presented in this article focuses on the potential for interpretation afforded by corpus-based lexicography, its other main aim being an evaluation of the results and the cost in terms of time. The computerised system elaborated for this purpose can be extended to other projects, particularly those focusing on medieval Galloromania.

**Keywords :** Old French, computer-based lexicography, textual corpus, dictionary writing systems, lemmatisation

**Mots clés :** ancien français, lexicographie électronique, corpus textuel, système de rédaction, lemmatisation

---

**Martin Glessgen**, Romanisches Seminar der Universität Zürich, Zürichbergstrasse 8, CH-8032 Zürich, e-Mail: glessgen@rom.uzh.ch

**Marguerite Dallas**, Romanisches Seminar der Universität Zürich, Zürichbergstrasse 8, CH-8032 Zürich, e-Mail: marguerite.dallas@uzh.ch

<https://doi.org/10.1515/lex-2019-0009>

# 1 Idée de départ : un dictionnaire électronique basé sur des textes documentaires

Notre réflexion autour d'une exploitation lexicologique des textes documentaires français médiévaux remonte à 1998, quand nous avons projeté de continuer l'entreprise des *Documents linguistiques de la France* de Jacques Monfrin (Glessgen 2001). Notre réflexion a débouché relativement vite sur l'idée de concevoir un « Nouveau Godefroy » (Glessgen 2003), plus philologique et ciblé sur des textes documentaires. Notre idée, aussi simple qu'exigeante, était (1) de réunir dans une base de données relationnelle l'édition d'un nombre important de documents pour (2) permettre ensuite la rédaction d'un dictionnaire thématique, consacré à un ensemble de traditions de discours encore parfaitement négligé en linguistique diachronique<sup>1</sup>.

Cette idée était issue de nos expériences antérieures. D'une part, nous avons développé une certaine pratique en lexicologie philologique dans nos deux thèses sur un inventaire d'actes en occitan (Glessgen 1989) et un traité médical en ancien italien et en latin (Glessgen 1996). D'autre part, nous avons vu dans la rédaction des articles du *Lessico Etimologico Italiano* (LEI) combien un dictionnaire philologique de la langue médiévale était indispensable à la réflexion diachronique et étymologique – à l'époque le TLIO n'existait pas et encore moins d'autres ressources comme, par exemple, la *Biblioteca Digitale del Sud* (BDSud).

Dans la lexicographie galloromane, la lacune essentielle concernait – et concerne toujours – le langage documentaire, présent essentiellement dans le dictionnaire de Godefroy (Gdf). Cet ouvrage, qui force le respect par ses dimensions, couvre tous les siècles du Moyen Âge et tous les genres textuels, qu'ils soient littéraires, documentaires ou, dans une moindre mesure, scientifiques, se nourrissant autant d'éditions que de documents originaux et réunissant en dix volumes quelque 70 000 lexèmes. Il comporte toutefois des imperfections flagrantes : les sources ne sont pas facilement identifiables – problème aujourd'hui résolu grâce à la bibliographie de J.-L. Ringenbach (GdfBibl)<sup>2</sup> – et les transcriptions des originaux ne sont pas toujours fiables,

<sup>1</sup> Une version synthétique du présent texte a été prononcée lors du 29<sup>e</sup> Congrès International de Linguistique et de Philologie Romanes à Copenhague, en juillet 2019. Nous remercions les participants à la discussion pour leurs observations et Sabine Tittel ainsi que Thomas Städtler pour leur relecture attentive du manuscrit. – Les chapitres 1, 2, 4 et 5 du présent texte ont été conçus par M. Glessgen, le chapitre 3 par M. Glessgen et M. Dallas, qui a également réalisé l'essentiel du travail pratique de l'intégration des mots des DocLing dans le DEAFpré. Le pronom personnel « nous » renvoie ainsi dans le premier cas à M. G., dans le deuxième cas parfois à M. G. et M. D., parfois à M. D. toute seule ; l'accord est par conséquent, selon les cas, au masc. ou au fém. sg. ou encore au pl.

<sup>2</sup> Avec ses plus de 18 000 entrées, celle-ci permet d'identifier – en combinaison avec la bibliographie du DEAF (DEAFBibl) – la quasi-totalité des sources citées dans Godefroy ; cf. le site de l'ATILF (<http://www.atilf.fr/BbgGdf/>).

ce qui a mené à de nombreux « mots fantômes »<sup>3</sup>. Par ailleurs, le travail définitoire reste rudimentaire, puisque Gdf ne donne que de simples équivalences en français moderne, souvent sans distinguer correctement les sens des exemples cités. Ce dernier point est d'autant plus regrettable que le FEW reprend normalement tels quels les sens donnés par Godefroy, ce qui en augmente la diffusion<sup>4</sup>.

Or, dans les domaines du vocabulaire littéraire (religieux et profane) et également scientifique du Godefroy, la matière traitée par lui a été reprise à travers les cent dernières années par les autres grands dictionnaires de l'ancienne langue, le Tobler-Lommatzsch (TL), le DEAF, le DMF et l'AND<sup>5</sup>. Il est vrai que le domaine des sciences médico-biologiques comporte toujours un potentiel considérable, notamment si l'on prend en considération les nombreux mots vernaculaires latinisés dans les textes latins ; mais l'essentiel de ce qui se trouve dans le Godefroy a été répertorié par ses successeurs, qui ont élargi les données et qui en ont considérablement amélioré le traitement. Il reste toutefois la part des ressources documentaires, que la lexicographie récente prend naturellement en considération, mais où le Godefroy reste toujours incontournable. Rappelons que son auteur a été fortement soutenu dans sa collecte par des archivistes et des habitués d'archives.

Dans l'idée de réaliser un répertoire philologiquement mieux fondé que le Godefroy, nous nous sommes par conséquent concentré d'emblée sur l'ensemble des textes documentaires. Notre projet s'est placé dès les débuts dans la logique d'une lexicographie électronique. C'est la voie prise par le TLIO et l'OVI tout comme par le CORDE ou encore par de nombreuses autres entreprises, souvent d'envergure. La particularité – exclusive – de notre approche réside toutefois dans le fait de permettre des corrections dans les textes qui sont à la base du dictionnaire électronique, alors que la lexicographie sur corpus repose partout ailleurs sur une base textuelle *ne varietur*. Notre idée est donc celle d'un lien vivant entre texte et dictionnaire, où ce dernier ne prévoit pas seulement des hyperliens vers le texte mais peut avoir un impact sur les choix éditoriaux.

3 Ce problème a été décrit dans les années 1980/90 par J.-P. Chambon et Cl. Buridant ; la base des mots fantômes de N. Steinfeld (Steinfeld 2005–) identifie actuellement *ca* 700 entrées de ce type ; cf. le site de l'ATILF (<http://www.atilf.fr/MotsFantomes/>).

4 Cf. toutefois l'argumentation nuancée de Chauveau (2003) quant à l'utilisation du Gdf par le FEW.

5 Cf. Carles *et al.* (2019, 129–133) pour une présentation du Godefroy, et plus généralement *ib.* (128–178) pour un aperçu évaluatif de la trajectoire de la lexicographie diachronique galloromane depuis la fin du 19<sup>e</sup> siècle.

## 2 Les étapes préparatoires : édition et programmation

Pour le premier volet de notre entreprise – le fondement philologique du futur dictionnaire –, nous avons constitué à travers les quinze dernières années la base de données textuelle des *Documents linguistiques galloromans* (DocLing), soutenu par l'École Nationale des Chartes, par les Universités de Strasbourg et de Zurich ainsi que par la FNS et aidé par de nombreux collègues, amis et élèves. Actuellement, cette base comporte 1 250 000 occurrences contenues dans près de 3 000 actes provenant de textes du territoire d'oïl ainsi que 700 000 occurrences dans des textes provenant du territoire francoprovençal et étant partiellement rédigés en français (cf. Glessgen 2015, 2018 et Carles/Glessgen 2019). Le nombre d'occurrences reste limité – la BFM par exemple comporte 4,5 millions d'occurrences, l'ОВI presque 24 millions et le CORDE non moins de 250 millions –, mais contrairement aux autres bases de données du français médiéval<sup>6</sup>, il s'agit ici d'éditions pour l'essentiel *princeps* de documents originaux, datés et même localisables avec précision. Nous avons déjà présenté à diverses reprises les principes d'édition « mixtes », qui combinent des éléments diplomatiques et interprétatifs, et l'idée de vues alternatives sur un même corpus textuel (cf. Glessgen 2003).

Voici un exemple de la version mixte, organisée d'après des divisions sémantico-syntagmatiques, ici au nombre de onze :

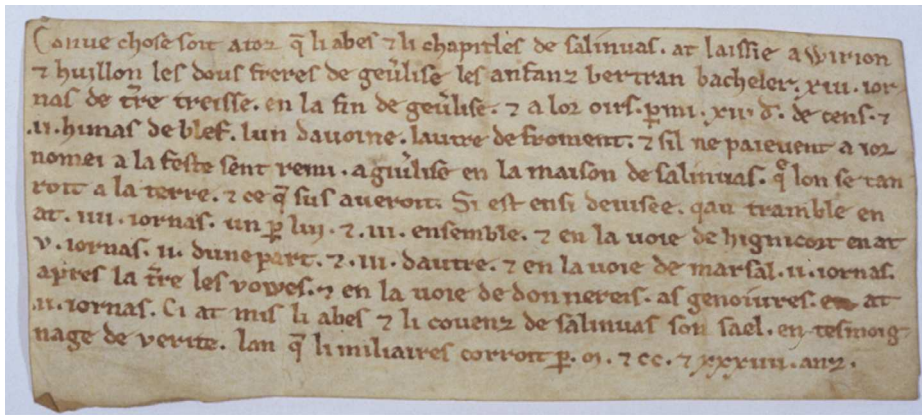
**1** Conue chose soit à-toz **2** que li abes et li chapitles de Salinvas· at  
laissé à Wirion <sup>\2</sup> et Huillon, les dous freres de Geverlise, les anfanz  
Bertran Bachelier, **3** /.XIII/. jor<sup>\3</sup>nas de terre treisse· en la fin de  
Geverlise· et à lor oirs· **4** parmi ·XIII/. deniers de cens· et <sup>\4</sup> ·II/.  
himas de blef· l'un d'avoine· l'autre de froment· **5** et s'il ne paievent à  
jor <sup>\5</sup> nomei à la feste sent Remi· à Giverlise, en la maison de  
Salinvas <sup>[i]</sup>. que l'on se tan<sup>\6</sup>roit à la terre· et ce que sus averoit· **6** Si  
est ensi devisee· q'au Tramble en <sup>\7</sup> at ·III/. jornas· un par lui <sup>[iii]</sup>. et  
·III/. ensemble· **7** et en la voie de Hignicort en at <sup>\8</sup> V/. jornas· II/.  
d'une part et ·III/. d'autre· **8** et en la voie de Marsal ·II/. jornas· <sup>\9</sup>  
après la terre les Vowes <sup>[iii]</sup>. **9** et en la voie de Donnereis· as  
Genoivres· en at <sup>\10</sup> ·II/. jornas· **10** Ci at mis li abes et li covenz de  
Salinvas son sael· en tesmoig<sup>\11</sup>nage de verité· **11** l'an que li miliaires  
corroit par ·M/. et CC/. et XXXIII/. anz·

Ill. 1. Exemple des principes d'édition critique des DocLing (chMM n° 2, 1234)

<sup>6</sup> Cf. Tittel, 2018, 455, n. 5 ; Carles et al. 2019, 165–167.

Et voici une vue diplomatique, qui reproduit de plus près la forme textuelle originelle :

- 1 Conue chose soit atoz *que* li abes *et* li chapitles de salinvas · at laissie a wirion
- 2 *et* huillon les dous freres de geverlise les anfanz bertran bachelier · XIII · jor
- 3 nas de terre treisse · en la fin de geverlise · *et* a lor oirs · *parmi* · XIII · deniers de
- 4 cens · *et*
- 5 ·II· himas de blef · lun davoine · lautre de froment · *et* sil ne paievent a jor
- 6 nomei a la feste sent remi · a giverlise en la maison de salinvas · *que* lon se tan
- 7 roit a la terre · *et* ce *que* sus averoit · Si est ensi devisee · gau Tramble en
- at · III · jor · un *par* lui · *et* · III · ensemble · *et* en la voie (...)



III. 2. Exemple d'une édition diplomatique, parallèle au ms. médiéval

Pour le volet lexicographique, nous avons élaboré, également à travers plusieurs années et avec un soutien informatique professionnel, un logiciel (*Phoenix2*) pour différents types d'interrogations et pour la lemmatisation des formes. Dans un premier temps, nous avions l'intention de mettre en place un dictionnaire électronique à part entière et nous en avons développé le prototype (cf. Glessgen/Alletsgruber 2010). Mais après mûre réflexion, nous avons renoncé à cette nouvelle base, qui se serait ajoutée aux multiples fragments lexicographiques qui ont vu le jour ces deux dernières décennies. Nous avons plutôt retenu pour l'exploitation lexicographique de nos données en 2011 l'idée d'intégrer le vocabulaire des documents jusqu'ici inédits dans la base de données du *Dictionnaire étymologique de l'ancien français* (DEAF).

Ce dictionnaire de référence pour l'époque du 9<sup>e</sup> au 14<sup>e</sup> siècle est achevé pour seulement une partie du vocabulaire (les lettres G-K et, partiellement E et F, = DEAFplus). Devant une impasse financière, ses matériaux englobant le restant de l'alphabet ont été saisis et ordonnés succinctement par les rédacteurs du DEAF créant ainsi

le « DEAFpré » en accès libre (Tittel 2010)<sup>7</sup>. Ce dernier est une réalisation provisoire puisque les occurrences individuelles n'ont pas pu être vérifiées lors de la rédaction, chaque fiche devant être traitée en huit secondes en moyenne<sup>8</sup>. Il est néanmoins précieux par la richesse de ses matériaux et par sa structuration étymologique raisonnée. Grâce à leur grande expérience en lexicographie, les rédacteurs ont su tirer le meilleur parti de cette mise en ligne « accélérée ». Le DEAFpré comporte par ailleurs la qualité de toute lexicographie électronique bien conçue d'être évolutive et de permettre des intégrations, des améliorations et des développements ultérieurs – qualité que nous avons souhaité exploiter par l'intégration des matériaux des DocLing<sup>9</sup>.

Si nous sommes toujours convaincu du bien-fondé de ce choix, nous avons dû nous confronter à la complexité et à la lenteur des processus de programmation, et ceci malgré un engagement inconditionnel de la part de notre amie Sabine Tittel et du soutien constant des programmeurs du DEAF, Conny Kühne et Marcus Husar. Ainsi, le prototype de l'intégration de nos données dans le DEAF était prêt dès 2013, quand nous l'avions présenté avec Sabine Tittel lors du *Congrès de Linguistique et Philologie Romanes* à Nancy. Nous avons toutefois renoncé, malgré le souhait des présidents de section, David Trotter et Andrea Bocchi, à la publication de cette présentation, parce que la programmation de l'interface entre les *Documents linguistiques* et le DEAFpré n'était que partiellement achevée. Elle s'est en effet avérée très épineuse, doublée d'une mise à jour intégrale du programme du DEAF à Heidelberg. L'interface entre les DocLing et le DEAFpré est pleinement fonctionnelle depuis la fin 2017 et la nouvelle version du DEAF est en ligne depuis avril 2019 (Tittel 2018, 2019). Comme nous l'avons déjà formulé ailleurs : l'informatique est à la fois une bénédiction et une malédiction pour la lexicographie diachronique (Glessgen 2018, 1258).

Nous souhaiterions présenter par la suite les différentes étapes parcourues pour réaliser l'intégration des lexèmes de la base textuelle des *Documents linguistiques* dans la base lexicographique du DEAFpré. Cela permet de cerner les implications méthodologiques d'un tel projet et d'identifier précisément les opérations nécessaires pour suivre cette voie qui nous semble probante et riche en perspectives. Par ailleurs cela permettra de quantifier l'apport lexicologique de nos données et d'évaluer les dimensions de ce que nous avons conçu vingt ans auparavant comme un « Nouveau Godefroy ».

7 Cf. Tittel (2018, 454, n. 2) : « The lemmata treated in the form of DEAFplus (letters D – K) will add up to approximately 9,000 in 2020 ; approximately 73,000 lemmata will remain as DEAFpré (the rest of the alphabet) for the time being ».

8 Cf. Tittel (2018, 454) : « DEAFpré is not a dictionary in its proper sense, but offers the complete raw material of the DEAF. »

9 Cf. aussi pour la genèse et les qualités du DEAF, Carles *et al.* (2019, 139–144).

### 3 Réalisation lexicographique

L'élaboration du lexique des *Documents linguistiques* dans un cadre lexicographique suppose cinq étapes bien distinctes et demandant un effort très variable :

- (1) la lemmatisation des lexèmes dans le cadre des *Documents linguistiques* (3.1)
- (2) l'importation des lemmes établis dans la base de données du DEAFpré (3.2)
- (3) l'intégration des formes lexicales dans les articles du DEAFpré (3.3)
- (4) l'intégration sémantique des lexèmes dans le DEAFpré (3.4)
- (5) la rédaction d'articles lexicologiques dans le cadre du DEAFpré (3.5)

Nous commenterons par la suite ces étapes individuellement<sup>10</sup>.

#### 3.1 La lemmatisation

La première étape de tout travail lexicologique consiste dans le rapprochement des variantes formelles d'un lexème sous un lemme donné. La lemmatisation des DocLing a commencé dès 2005 avec une version antérieure de *Phoenix*, programmée alors sous Tustep (Glessgen/Kopp 2005). Les premiers corpus disponibles, la Meuse et la Meurthe-et-Moselle, ont été lemmatisés par Anne-Christelle Matthey-Bovet, Julia Alletsgruber et Dumitru Kihăi. Le travail s'est élargi au fur et à mesure à travers les années, comportant plusieurs révisions, notamment par notre regrettée élève et amie Claire Vachon, et un nombre considérable de corrections dans les éditions elles-mêmes, coordonnées dernièrement par Tania Paciaroni et Marguerite Dallas. Nous ne pouvons que donner raison à Frankwalt Möhren, qui a toujours insisté sur le fait que l'élaboration d'un glossaire permet une amélioration considérable de l'édition. Malgré l'effort de programmation considérable que cela a entraîné, la possibilité de corriger les éditions suite aux résultats de l'analyse lexicologique nous semble toujours essentiel.

Quant à la forme des lemmes, celle-ci n'a pas une importance capitale, puisque le but premier d'une lemmatisation est simplement de constituer des ensembles lexicaux identifiables. Mais l'attribution des lemmes demande néanmoins une réflexion certaine et une logique cohérente<sup>11</sup>. Dans un premier temps, nous avons choisi des lemmes s'approchant au mieux des formes en français moderne pour faciliter l'accès aux mots anciens. Quand nous avons décidé d'intégrer notre corpus dans le DEAF,

<sup>10</sup> Notre réflexion est complémentaire de celle de notre partenaire de projet Sabine Tittel qui explique la même procédure d'un point de vue technique et informatique (Tittel 2018, 464–456 ; ce texte repose sur le chapitre 9.4.6 de la thèse d'habilitation Tittel 2019). Nous insisterons ici sur les aspects lexicologiques et lexicographiques de notre projet commun.

<sup>11</sup> Cf. par exemple Möhren 2015, 422–424.



nous avons dû remplacer nos lemmes par ceux du DEAF, inspirés du Tobler-Lommatzsch et ciblant l'état phonique des mots aux 12<sup>e</sup>/13<sup>e</sup> siècles. L'adaptation s'est faite à travers une liste de concordance, permettant de nouveau un certain nombre de corrections. Nous verrons qu'il reste des cas problématiques, puisque les formes héréditaires et savantes d'une même base se placent souvent dans la logique d'un continuum plus que d'une opposition (cf. *infra* 3.2 sous (2)).

Notre module de lemmatisation suppose un travail manuel, aidé par l'informatique. Les variantes formelles sont placées dans l'ordre alphabétique à droite, accompagnées de leur nombre. En cochant une forme, les occurrences apparaissent avec le contexte dans la fenêtre centrale, où elles peuvent être organisées par forme, date ou corpus. Cela permet de retenir assez rapidement les occurrences appartenant à un même lemme, sachant qu'il est possible de voir le contexte intégral des documents en cliquant sur les formes. Voici l'exemple de *baillier* :

The screenshot displays the Phoenix2 software interface for lemmatization. The main window is titled 'Assignment' and shows 'Assign Occurrences' and 'Restrictions' tabs. Below these, a status bar indicates 'Assignment successful: 14 Occurrences assigned to Lang «fr»'. The interface is divided into several sections:

- Assign Lemma to Selected Occurrences:** This section includes a dropdown for 'Existing Lemma' (currently set to 'baillier, baillier [c]') and buttons for 'Assign Existing Lemma', 'Cancel', and 'Clean Empty Lemmata'. It also has a 'New Lemma (see reference)' section with fields for 'Lemma (identifier)', 'Main identifier', 'Type' (set to 'c'), and 'PoS', along with 'Assign New Lemma' and 'Cancel' buttons.
- Table of Occurrences:** A table with columns: 'Clef', 'Date', 'Red.', 'Div', and 'Context'. It lists various occurrences of the word 'baillier' and its variants, such as 'baillies', 'baillier', 'baillera', 'baillierai', 'baillierent', 'baillierens', 'baillierent', 'baillieront', 'baillieron', and 'baillierons'. Each entry is preceded by a checkbox.
- Browse Panel:** Located on the right, it shows a list of lemmata with their counts in parentheses: 'baillies (5)', 'baillier (3)', 'baillier (19)', 'baillier (9)', 'baillera (4)', 'baillierai (1)', 'baillierent (2)', 'baillierens (3)', 'baillierent (1)', 'baillierole (1)', 'baillieront (4)', 'baillieront (1)', 'baillieron (1)', and 'baillierons (1)'. There are also checkboxes for 'select all' and 'show langified occ.'.
- Query (REGEX):** A field at the top right with the value '\*ba' and an 'OK' button.

### III. 3. Module de lemmatisation du logiciel *Phoenix2*

Le logiciel permet aussi d'appréhender les occurrences une fois réunies sous un lemme. L'exemple de *angalt* 'droit qu'on perçoit sur le vin' laisse entrevoir les variantes qui peuvent se trouver en un seul document (*angalt*, *angals*, *engaul* et *angaul* ; les parenthèses indiquent qu'il s'agit de lettres abrégées dans le manuscrit) :

**Phoenix2**  
Web-Based Annotation

Version 0.7 Alpha Build 2

Assignment Assign Occurrences Restrictions

☐ All texts | ☒ Selected texts only (see Restrictions)

Occurrences +Lemma +Morph +Graph +Lang Export Clear Results

	CstF	Date	Red.	Div	Context
<input type="checkbox"/>	doJuBe168	1339	-	117	ent de l - a(n)galt de Porreintruy . .XXXV. lb . .XVI. s . et .X. d . Item
<input type="checkbox"/>	doJuBe168	1339	-	121	aut de l - a(n)galt dou pahis . .VIII. lb . et .V. s . que W . diz Paniaste
<input type="checkbox"/>	doJuBe168	1339	-	131	valet li a(n)gals dou pahis .XXX. lb . Item . li banz devant l - aule .
<input type="checkbox"/>	doJuBe168	1339	-	137	oma de l - angalt de Porreintruy : . Vixx . lb . et .XXX. s . . V. . Soma so
<input type="checkbox"/>	doJuBe168	1339	-	139	Valet li angals de Porraintruy cent et .XVII. lb . .VII. s . et .VI. d .
<input type="checkbox"/>	doJuBe168	1339	-	166	son de l - engaul dou pais , dou temps Bourquart dit Chamon . Item . .VIII
<input type="checkbox"/>	doJuBe168	1339	-	170	n , de l - engaul dou pais . Item . .XXVI. s . des diz harcheret et Hestev
<input type="checkbox"/>	doJuBe168	1339	-	171	n , de l - engaul dou pais , le merdi après lai feste de lai Nativeteyl No
<input type="checkbox"/>	doJuBe168	1339	-	172	t pour l - engaul dou pais et .XII. s . le sambbadi après lai Saint Luc .
<input type="checkbox"/>	doJuBe168	1339	-	174	z pour l - engaul dou pais , le merdi après lai Saint Andree , C . s . Item
<input type="checkbox"/>	doJuBe168	1339	-	177	z , de l - e(n)gaul dou pais . Item . .VII. s . de Hestevenin et de l - ar
<input type="checkbox"/>	doJuBe168	1339	-	178	t pour l - engaul dou pais , le juefdi après l - an - nef . Item , ont pah
<input type="checkbox"/>	doJuBe168	1339	-	179	s , de l - e(n)gaul dou pais . Item , de Chiotat . .VI. s . //, de plusieurs
<input type="checkbox"/>	doJuBe168	1339	-	240	Item , li angaul de l - aifuert . .XIV. lb . vriez bailloys a bon us dou p

Browse Lemmata Types

Query (REGEX):  
angalt OK

Category: ☐ p ☐ c ☐ b ☒ cp ☒ d

Lang: ☒ no lang ☒ fr ☐ frfr ☐ frpr ☒ lat

☒ show langified occ. ☐ only zero occ.

☒ ongueul , angalt(14)

#### III. 4. Variantes et occurrences d'un lemme donné dans le logiciel *Phoenix2*

Suite à l'intégration récente des corpus du Jura bernois, de la Côte d'Or et de la Moselle, lemmatisés par Alessandra Bossone et Marco Robecchi, nous avons nouvellement revu l'état de la lemmatisation et nous avons réuni une liste intégrale des lemmes de nos corpus.

D'un point de vue quantitatif, la récolte lexicale des *Documents linguistiques* atteint un volume certain, mais elle reste malgré tout circonscrite. L'intégralité des 3 000 actes oïliques du 13<sup>e</sup> et du début du 14<sup>e</sup> siècle comporte environ 4 600 formes lexicales lemmatisées. Il est vrai que le seul traité de médecine aviaire du *Moamin* que nous avons édité en son temps comporte – les deux manuscrits confondus – presque le même nombre de formes lexicales. L'idée que les actes sont un genre stéréotypé n'est donc pas pleinement erronée. À l'avantage des DocLing il faut toutefois dire que la fréquence des occurrences sous un même lemme est souvent considérable, tout comme sa variance formelle. Par ailleurs, les formes sont souvent polysèmes ou intégrées dans des syntagmes qui permettent de mieux cerner leurs contextes d'usage habituels. Néanmoins, nous avons constaté déjà relativement tôt que les aspects scriptologiques, syntaxiques et onomastiques des actes renferment un potentiel linguistique peut-être plus important que celui de la nomenclature lexicale.

Ce bref aperçu montre que la première étape du travail lexicographique a supposé un effort considérable et une interaction constante entre la réflexion lexicologique et la philologie éditoriale. La lemmatisation nous a ainsi occupé somme toute une bonne décennie, sachant qu'elle a contribué sensiblement à l'amélioration des éditions et que nous avons travaillé en même temps sur différents chantiers du projet des DocLing. Dans la logique d'une mise en valeur rapide des projets de recherche, propre à l'époque actuelle, cet investissement représente un enjeu problématique puisque le travail réalisé par nos collaborateurs et nous-même est resté pendant longtemps non

apparent. Ce n'est que par l'intégration dans le DEAFpré que les résultats d'un travail constant et décennal seront devenus visibles.

### 3.2 L'importation des lemmes dans la base de données du DEAFpré

Une fois la lemmatisation achevée, la véritable intégration dans le DEAFpré a pu commencer. La première étape correspondait alors à l'harmonisation de nos lemmes avec ceux du DEAFpré. Ces derniers sont en effet organisés de manière hiérarchique et distinguent des lemmes principaux, correspondant pour l'essentiel à une base latine ou protoromane donnée, et des lemmes subordonnés, correspondant *cum grano salis* à des dérivés d'époque romane<sup>12</sup>. Il s'agit donc de lemmes bipartites.

Les lemmes que nous avons retenus pour les DocLing représentent la dimension des lemmes subordonnés, donc les lexèmes individuels en ancien français. Il a ainsi fallu organiser ces lemmes « simples » dans une logique étymologique d'après l'époque de leur formation, en regroupant les lexèmes ayant une même base latine ou germanique antérieure à *ca* 700<sup>13</sup>.

Cette opération supposait donc concrètement une comparaison des lemmes simples des DocLing avec les lemmes doubles du DEAF et, en cas d'ambiguïté, un choix pour rapprocher correctement les entités à l'étude. Les choix déjà établis dans le DEAFpré ont servi de modèle, ce qui a facilité le travail et augmenté la sécurité de décision.

Voici un exemple de l'interface programmée par l'équipe du DEAF pour l'identification des lemmes « principaux » à partir des lemmes retenus dans les DocLing :

---

<sup>12</sup> Cf. Tittel (2018, 454) : « The DEAF organizes the Old French lexicon in word families to show the etymological relations between single lexemes. The *main-lemma* of a dictionary entry is the lexeme that is developed or borrowed from a Latin, Greek, etc., origin. The derivations from this lexeme are the *sub-lemmata*. » – C'est également la logique que nous avons retenue pour le *Dictionnaire d'ancien gascon* électronique (DAGél). Le DEAFpré est moins rigoureux sur ce point et suit parfois le FEW, qui regroupe sous un étymon souvent ses dérivés latins attestés – E. Buchi a réuni non moins de 10 000 exemples de ces « étymons cachés » (1996, 405–563) – et toujours les dérivés non attestés à l'époque antique (= les formations protoromanes).

<sup>13</sup> S'ajoute le cas relativement rare des emprunts au latin médiéval, notamment dans le vocabulaire scientifique ou juridique.

Attribution of DocLing lemma *parent* to DEAF lemma

## Selection DEAF lemma

Search with wildcard character: "%" matches any number of characters, "\_" matches a single character.

Showing 1 to 3 of 3 << < 1 > >>

Main lemma	Sublemma	Assign
en2	parent	<a href="#">Assign</a>
parent	parent	<a href="#">Assign</a>
per	parent	<a href="#">Assign</a>

[Cancel](#)

## Occurrences

&lt;&lt; &lt; 1 &gt; &gt;&gt;

## Word in context / Metadata occurrence

Ceu est la maniere et li estableissement de la commune pais de Mez //. Qui homme ocira dedenz la pais //, et tesmognez en sera //. sei **parent** ocirunt lui et sexante livres i panra la justice /. de sun moule /. ou de sun heritage /. si li moules n' i estoit //. et toz li remanz avera pais /. et li ami /. Et si li ami n' avoient la force //. et li lo manife [↗](#)

1215/01/01 — chMo001 2 Red.: - Scripta: - Text type: -

## III. 5. Module d'identification des lemmes principaux du DEAFpré à partir des lemmes (secondaires) des DocLing

Le lemme *parent* des DocLing est rapproché de manière automatique des trois lemmes principaux qui, dans le DEAFpré, connaissent ce lemme subordonné : les bases adverbiale *en2* (< INDE) et adjectivale *per* (< PAR) ainsi que la base nominale *parent* (< PARENS). Une vérification dans le DEAFpré permet de préciser le sens des trois lexèmes homographes, l'adverbe (peu fréquent) 'par ci, par là, dans les environs ; dorénavant', l'adjectif (rarissime) 'égal, pareil' et le substantif (très usuel) 'personne qui a un lien de parenté avec quelqu'un, qui est membre de la même famille'. Le contexte des DocLing reproduit dans l'interface même permet immédiatement l'identification du mot avec ce dernier, sachant que le contexte concret n'est pas banal : « celui qui tuera un homme à l'intérieur de l'espace de validité de la paix territoriale [*pais*] et si le fait est établi par témoignage, ses *parents* [= les parents de la victime] le tueront [à son tour] et la justice prendra 60 livres de ses biens [= les biens de l'assassin ; *moule* est une forme régionale de *meuble*] ou de son héritage ».

Notons que dans cette étape, on voit encore une fois le contexte des mots dans les DocLing, ce qui permet de prendre en considération la dimension du sens au moment du rattachement à un lemme donné du DEAF. Grâce à ce contexte, il est également possible, dans le cas concret, de rapprocher l'occurrence du substantif *parent* sans passer par une vérification du sens des trois lemmes principaux dans le DEAFpré. Un rattachement donné peut de toute manière être corrigé à tout moment.

Notons par ailleurs un point méthodologique particulier dans la gestion des lemmes : pour éviter des malentendus, dans le cas d'une forme non dérivée, le lemme principal du DEAF correspond en même temps à un lemme subordonné identique.

Voici l'exemple de *testament*, où les entrées apparaissent sous la forme suivante dans le système du DEAFpré :

Lemma 1 bis 5 von 5

	<input type="text" value="testament"/>	<input type="text"/>	<input type="button" value="Filter"/>	<input type="button" value="Clear"/>
Lemma	Main lemma	Sublemma	Short article	Unassoc.
<a href="#">Edit</a>	testament	testamenteor	<a href="#">Modify</a>	108
<a href="#">Edit</a>	testament	testament	<a href="#">Modify</a>	113

### III. 6. Relevé de deux lemmes doubles dans le système de rédaction du DEAFpré

La dernière étape de l'intégration correspond à une mise à jour des lemmes des DocLing, à partir du rapprochement opéré dans la base du DEAFpré. Si donc au début, les DocLing ne connaissaient que le lemme simple « testament » et « testamenteor », suite à cette mise à jour, le mot de base sera lemmatisé dans notre logiciel *Phoenix2* comme « testament, testament » (avec 113 occurrences), à l'instar du DEAFpré, et le dérivé *testamenteor* comme « testament, testamenteor » (108 occ.) :

Browse

Lemmata

Types

Query (REGEX):

Category: ☒ p ☒ l ☒ c ☒ lp ☒

cp ☒ cl ☒ lat

Lang: ☒ no lang ☒ fr ☒ frfrpr ☒

frpr ☒ lat

☒ show langified occ. ☐ only zero occ.

☐ select all

☐ testament, testament (113)

☐ testament, testamenteor (108)

### III. 7. Relevé de ces mêmes deux lemmes doubles dans *Phoenix*

La comparaison systématique des lemmes principaux et secondaires a encore mené à un certain nombre de corrections parmi les lemmes de DocLing et également parmi ceux du DEAFpré. Comme nous l'avons déjà mentionné (cf. *supra* 2), ce dernier représente une collection de matériaux provenant du fichier manuel du DEAF – des « Matériaux bruts non vérifiés », comme l'indique le site –, et non une matière élaborée dans les règles de l'art.

En tout, nous avons importé environ 4 300 lemmes, en excluant quelque 200 mots à cause de leur haute fréquence. Cela concerne notamment une série de mots grammaticaux avec plus de 100 occurrences dans notre corpus, par exemple :

enz	101 occ.
nïent	106 occ.
jadis	107 occ.
soi1	814 occ.
tant	825 occ.
mon1	841 occ.
devant	971 occ.

Dans certains cas, l'équipe du DEAF nous a conseillé de renoncer aussi à des mots lexicaux à haute fréquence, comme par exemple :

vendre	909 occ.
aumosne	914 occ.
nomer	959 occ.
grace	979 occ.
chevalier	985 occ.

Dans un premier temps, nous avons également exclu *tous* les mots avec plus de 1000 occurrences à cause de la lourdeur de leur gestion. Cela concerne encore 36 lexèmes comme *tesmoignage* (1720 occ.) ou *bois* (1035 occ.).

Pour une meilleure gestion des données, nous avons établi une liste intégrale des correspondances retenues. En voici le début, dans l'ordre alphabétique des lemmes subordonnés :

lemme principal	lemme subordonné
aatir	aatine f.
baissier	abaissier v.
ban	abandon m.
ban	abandoner v.
abatre	abatre v.
abeesse	abeesse f.
abeie	abeie f.

soing	abesoignier v.
abevrer	abevrer v.
abit	abit m.
abitacion	abitacion f.
abiter	abitant p. prés. s.
abiter	abiter v.
abondance	abondance f.
abonder	abondement m.

Dans l'attribution des lemmes, nous avons dû résoudre de nombreuses questions lexicologiques de détail. Les cas de figure les plus récurrents sont les suivants :

(1) Dans les cas où les articles du DEAFpré suivent le FEW dans le regroupement des dérivés déjà latins avec le mot de base (cf. *supra* n. 12), nous avons séparé les deux ensembles.


Un exemple : le cognat *circonvencion* emprunté au latin classique CIRCONVENTIO, se trouve sous le verbe CIRCUMVENIRE dans l'article du FEW (2, 707b) et avait donc été placé dans le DEAFpré sous le verbe *circonvenir* :

**CIRCONVENIR** v.tr.

[TL; TL 2,440,24; TLF; LarLFr S.743 Bd.2. – DiekampSyn §43, §2; DiekampSyn; MessnerBers 72; ZfSL 84, p.336.]

(*circumvenir* [AalmaR 1865])

◆ “tromper qn, abuser qn” ([AalmaR 1865], DiekampSyn §43, §2; DiekampSyn; MessnerBers 72; ZfSL 84, p.336, TL; TL 2,440,24; TLF; LarLFr S.743 Bd.2)

Cf. DMF [CIRCONVENIR](#) .


---

● **circonvencion** f.

[TL IX, 97c; Gdf Suppl.9, 97c; AND 99a; MED 2,280a. – ActesMfr<sup>4</sup> p.111; Bartsch p165; DiekampSyn §16; Morlet 85; OrelliBibel; ZrP Bd.9x(1978) S.178.]

(*circonvencion* (sigles à datations multiples:) SchwanBehrens 3; SchwanBehrens p. 134b, *circumvencion* CartEngMarF p.152, 184, 205 etc.; [AalmaR 1866])

◆ “action de circonvenir par des artifices” (CartEngMarF p.152, 184, 205 etc.; [AalmaR 1866]; (sigles à datations multiples:) SchwanBehrens 3; SchwanBehrens p. 134b, ActesMfr<sup>4</sup> p.111; Bartsch p165; DiekampSyn §16; Morlet 85; OrelliBibel; ZrP Bd.9x(1978) S.178, TL IX, 97c; Gdf Suppl.9, 97c; AND 99a; MED 2,280a)

Cf. DMF [CIRCONVENTION](#) .

### III. 8. Ancien article *circonvenir* du DEAFpré

Selon notre logique rigoureusement étymologique, nous avons attribué à *circonvencion* le statut de lemme principal lors de l'importation de nos formes dans le DEAFpré.

(2) Dans d'autres cas, le DEAFpré a placé sous deux lemmes différents les variantes grapho-phonétiques reflétant une même base étymologique. Cela peut être une simple erreur, facile à corriger, mais cela peut également demander une réflexion plus poussée, notamment lors de la séparation entre formes à évolution phonétique héréditaire et emprunts savants. Comme nous l'avons déjà dit, le cas idéal d'une opposition binaire ne se réalise que rarement. Parfois, la gamme des formes intermédiaires présentées dans les DocLing rend un classement binaire impossible. Ainsi, le DEAFpré distingue correctement les lemmes *octobre*, comme lemme principal et mot savant, et *uitembre*, comme lemme subordonné héréditaire de *uit*. Dans la charte de la Meuse n° 42 nous trouvons toutefois la forme *ottembre*, qui montre une interférence entre les deux trajectoires. Nous l'avons placée sous *uitembre*, en considérant ce lexème comme déterminant, en plaçant un renvoi sous *octobre*.

(3) De manière plus concrète, la gestion des homonymes s'avère souvent délicate, d'autant plus que la tradition lexicographique n'est pas homogène dans la gestion des formes. Un exemple : la forme *lait* se trouve dans le DEAFpré sous les lemmes principaux *lait1* 'laid' (< abfrq. \**laip*, FEW 16,439a ; Gdf *lait4* ; DMF *lait2*), *lait2* 'lait' (< LAC, FEW 5, 110a ; Gdf *lait5* ; GdfC *lait2* ; DMF *lait1*), *lai2* 'lac' (< LACUS, FEW 5, 125b ; DMF *lai3*) ou encore *lai3* 'laïque' (< LAICUS, FEW 5, 131b ; Gdf/DMF *lai1*). Même si les sémantismes et les catégories grammaticales sont nettement distincts, le grand nombre d'attestations homographes a mené dans le DEAFpré à des erreurs notamment dans la distinction entre *lait1* 'laid' et *lait2* 'lait'. La catégorisation divergente dans la tradition complique le traitement de ces formes ultérieurement.

C'est là un cas extrême. Un problème plus simple, mais plus délicat d'un point de vue sémantique est la différenciation des occurrences de *devis* d'après les étymons \*DIVISARE (*devis* n.m. 'division, partage ; désir, intention', FEW 3, 109a) et DIVIDERE (*devis* part.pass. ou adj. de *devire* 'diviser') qui avaient été réunies sous un même lemme dans le DEAFpré.

En ligne générale, la superposition d'homonymes n'est pas rare dans le DEAFpré, toujours générées par la rapidité de sa réalisation (cf. *supra* 2). La séparation de ces formes a par conséquent demandé une attention constante lors de l'intégration de nos lemmes et il est important de savoir que les homonymes doivent être appréhendés avec précaution dans la base du DEAFpré.

(4) Enfin, la gestion du tréma s'est avérée une source constante de perturbations. Le DEAF place les trémas d'une manière différente du TL et nos éditions peuvent être en accord avec l'un ou l'autre des dictionnaires ou encore prendre une voie différente. Par exemple le lemme *diemaine* (< DOMINICUS, FEW 3, 129a) apparaît sans tréma dans le DEAFpré, tandis que les éditions des DocLing ont suivi le TL *diemaine* : le choix du DEAF est généralement plus convaincant – comme c'est le



cas pour *diemaine* –, mais la divergence nous oblige toujours à une intervention de part ou d'autre.

L'importation des quelque 4 300 lemmes et des actuellement 120 000 occurrences dans le DEAFpré nous a occupé à travers l'année académique 2018/19, avec un soutien constant de Sabine Tittel et Thomas Städtler. La procédure en tant que telle est rapide, mais les corrections demandent un temps certain puisqu'elles supposent une évaluation lexicologique plus ou moins poussée. La programmation de l'interface entre les DocLing et le DEAFpré en revanche s'est échelonnée à travers près de cinq ans et elle a demandé un certain nombre de réflexions méthodologiques. Ainsi, nous avons décidé sur les conseils de S. Tittel d'importer les occurrences des DocLing avec le contexte des divisions en question (cf. *supra* 2, l'exemple de l'acte MM n° 2) dans la base de données du DEAFpré. Nous avons en revanche rejeté l'idée initiale de faire apparaître les contextes des DocLing de manière dynamique (« on the fly »), au moment de la consultation, choix retenu en revanche pour le site internet des DocLing. Cela garantit une plus grande stabilité et sécurité philologique au DEAFpré, même si cela oblige à des mises à jours automatisées assez régulières pour intégrer dans le DEAFpré les éventuels changements dans la base textuelle des DocLing.

D'un point de vue lexicologique, cette importation fournit un apport concret et un apport méthodologique. D'un point de vue concret, elle a préparé l'intégration du vocabulaire des DocLing dans le DEAFpré (cf. *infra* 3.3 et 3.4) et, par ses corrections diverses, elle a apporté une contribution à l'élaboration du DEAFpré.

Au delà des mots traités par nous – environ 6 % des 73 000 entrées du DEAFpré (cf. *supra* n. 7) –, la réflexion menée sur les lemmes a permis d'identifier les points qui méritent une attention particulière dans la lemmatisation du DEAFpré et qui pourront aider ses utilisateurs.

### 3.3 L'intégration des formes lexicales dans les articles du DEAFpré

Une fois la lemmatisation et l'importation achevées, l'intégration des formes lexicales des DocLing dans le DEAFpré devient une opération presque mécanique et assez rapide. Il faut compter moins d'une minute par lemme.

Il faut en effet distinguer, pour l'intégration des lemmes, deux étapes, une purement formelle, une sémantique, selon la structure du DEAF, qui distingue toujours rigoureusement l'inventaire des formes de celui des sens et syntagmes. Nous reviendrons sur la deuxième étape, plus exigeante, dans la prochaine section (3.4). La première étape en revanche ne suppose qu'une intervention simple à l'intérieur du système rédactionnel du DEAFpré pour que les formes des DocLing soient reproduites dans le cadre des « articles courts » du DEAFpré.

Après l'importation, les occurrences des Docling paraissent dans le volet formel en un bloc à la suite des occurrences du DEAFpré. Dans le cas le plus fréquent où le lexème des DocLing connaît déjà un article court dans le DEAFpré, il suffit de cliquer sur le bouton « Sort by word on fiche / sort chronologically » au-dessus de la liste des formes intégrées<sup>14</sup>:

palais2 paladin palasin palasine palasio palatin palatine palantien palatinois

Dictionary block title block **Written forms** Senses Comment

All primary fiches for lemma

Manual sorting

Sort by word on fiche / sort chronologically << < 1 2 3 4 5 6 > >>

Fiche entry	Part of speech	Date	Siglum	Scripta	Reference	Variants	Definition	Rest	Order	Written form	Hide?
0 palais		ca. 1223	DolopL	lorr.				1053, 6363 palais.			<input type="checkbox"/>
1 palais		14es.	GallenD					m. palais, 3552			<input type="checkbox"/>
2 palais	adj.	ca. 1192	PrêtreJeand'yG	agn.	609		adj. du palais	Yale 777			<input type="checkbox"/>
3 palais		déb. 13es.	PercDidd'eR		app. A 6		Palast				<input type="checkbox"/>
4 palais		fin 11es.	AlexisH	norm.	1168						<input type="checkbox"/>

III. 9. Répertoire des occurrences pour un lemme donné (*palais2*) dans le système de rédaction du DEAFpré

Par cette opération, les formes des DocLing se placent dans la base (et ensuite sur le site) du DEAFpré en leur lieu alphabétique et chronologique. Voici l'exemple de *palais2* où l'on voit en troisième position une forme de la Saône-et-Loire (*pailois*).

<sup>14</sup> L'opération se complique quand il s'agit d'un nouveau lemme principal ou d'un nouveau lemme subordonné ; il faut alors ajouter la mention de la langue et de la catégorie grammaticale. Il faut également intervenir dans le système rédactionnel quand un lemme subordonné existe dans le DEAFpré, mais ne comporte pas d'occurrences (et n'est donc pas visible sur le site).

palais	paladin	palasin	palasine	palasio	palatin	palatine	palatien	palatinois
Dictionary block	title block	Written forms	Senses	Comment				

All primary fiches for lemma

Manual sorting

Sort by word on fiche / sort chronologically

Fiche entry	Part of speech	Date	Siglum	Scripta	Reference	Variants	Definition	Rest	Order	Written form	Hide?
0	palais	ca. 1223	DolopL	lor.				1053, 6363 palais.			<input type="checkbox"/>
1	palais	14es	GalerD					m. palais, 3552			<input type="checkbox"/>
2	palais	1316/00/00	chSL 068	bourg	23						<input type="checkbox"/>
3	palais	adj.	ca. 1192	PrêtreJeanDyG	agn.	609	adj. du palais	Yale 777			<input type="checkbox"/>
4	palais	delib. 13es	PercDiddéR	app. A 6			Palais				<input type="checkbox"/>
5	palais	fin 11es	AlexisH	norm.	1168						<input type="checkbox"/>
6	palais	fin 11es	AlexisP	norm.	O 18c, p. 73, 107						<input type="checkbox"/>
7	palais	ca. 1100	RoB	af.				s. palais [page = tome 2]			<input type="checkbox"/>
8	palais	sm.	1erq. 12es	BrendanW	agn.	S. 202, 279;	-elz (273),	palace			<input type="checkbox"/>
9	palais	2eq. 12es	GrantMarS2	norm.	82c, 96b		Palais				<input type="checkbox"/>
10	palais	2et. 12es	CoutLouisL1				palais, salle du palais	Gloss. Levy			<input type="checkbox"/>
11	palais	m. II	ca. 1160	EneasS1	norm.	460, 7466	palais				<input type="checkbox"/>
12	palais	ca. 1160	FloresK	af.	828, 832, ...		Palas, Hauptgelände d. Ritterburg...				<input type="checkbox"/>

### III. 10. Vue sur le même répertoire après l'intégration graphématique d'une occurrence des DocLing

Notons en passant que l'ordre des formes dans le DEAFpré est purement automatique ; celles-ci sont présentées d'abord d'après l'alphabet, puis – pour une même forme – d'après les dates, ce qui fait que la première attestation (ici : *palais*, fin 11<sup>e</sup> s., Alexis) se place souvent au milieu de l'inventaire des formes et non pas au début.

Pour revenir aux DocLing : une fois coché le bouton « sort by word on fiche / sort chronologically », l'article doit être nouvellement exporté à partir du système rédactionnel sur le site en accès libre. Les occurrences des DocLing y apparaissent alors en rouge et sont placées en leur lieu – organisées par forme et date – dans l'inventaire des formes du DEAFpré. Voici l'exemple de *ligement* :

#### • **ligement** adv.

[Gdf IV 781a; Gdf s. *ligement*; AND 388a; AND 382a; DCCarp 253c; v. gloss.; FEW s. *linge*, adj.; DupinLaboulaye. – ContPercIR T 12338; CorleyCont<sup>2</sup> 28291; Morlet 273; MéIMel 221; Runk p. 108; SchelerJPreis 30901 etc.; [sigle].]

(lie[g]emant chCOR013 3, *liegement* ChMM019 7, *ligemant* DolopL; ChMM062 4; chCOR013 5; chCOR032 6; ChMa132 4; ChMM219 6; ChMa205 10, *ligemant*, adv. (sigles à datations multiples:) SchwanBehrens p. 140b, *ligement* BenTroieC; BenDucM I, p. 486, v. 1172; AlexParA (ß 87) 1403 var. t. 5; IpH 7633; CoucyChansL XXX, 26; GaceBruléD \*6,20; Gloss.; MerlinM; DurmG gloss.; MoniotArtD v. gloss.; GuillVinM II, 14; ChansArtB XVIII, 189; CoutVerdun<sup>1</sup> M Anm. S. 152 zu Vers 22,60; DocHMarneG 124, 6; DocHainR 8; 2; ChMe012 5; chMM014 4; chMe025 3; ChMM040 9; ChMM053 4; ChMM058 15; ChMM061 4; BestAmOctT 10, 508; v. gloss.; HistFécL 2249; JeuxPartL; ChMM078 3; ChMM087 4; ChMM087 5; chCOR032 5; ChMM091 3; ChMM153 5; ChMM165 4; ChMM168 7; ChHM124 8; ChMM214 3; GilebBernH S. 80 v. 4 s. Ag. W; GilebBernF; ChMM223 10; ChMM238 3; ChMM238 4; ChMM238 4; ChMM246 4; ChMM279 10; ChMM285 3; chCOR184 7; chMe237 4; EtSLouisV II, 396; chCOR032a 12; chN 008 9; docJuBe017 8; docJuBe017 8; chN 023 3; TombChartr1/2/3S; [ProstInv 519; DeschQ VI, 126 etc.]; (sigles à datations multiples:) RaynMotets I, p. 52, etc.; StimmingMot, *ligement*, adv. BenDucF 13845, 29241; GaceBruléH; PrêtreJeanDyG 606; GodinM 8505, 11428, 16864; TournAntW 1688; AcartH 1802, *ligemant* ChMM233 3, *ligement* Bueve3S 14484(s. Anm.); ClefD 2510, *ligement* chMM023 19; chMM023 20; ChMM058 3)

### III. 11. Occurrences des DocLing intégrées dans le répertoire des formes du DEAFpré en ligne

Les formes provenant des DocLing ne représentent alors pas de simples renvois mais des hyperliens : en cliquant sur une référence donnée, on obtient le contexte de la division sémantique en question ainsi que les indications essentielles de l'ancrage diasystématique (date, lieu d'écriture et genre textuel) :

### • **ligement** adv.

[Gdf IV 781a; Gdf s. *ligement*; AND 388a; AND 382a; DCCarp 253c; v. gloss.; FEW s. *linge*, adj.; DupinLaboulaye. – ContPerc1R T 12338; CorleyCont<sup>2</sup> 28291; Morlet 273; MéiMel 221; Runk p. 108; SchelerJPreis 30901 etc.; [sigle].]

(*lie*gjen  
ChMM2  
BenDuc  
MerlinM  
22,60; D  
15; ChM  
chCOR0  
W; Gilel  
chCOR1  
TombCh  
Stimmin  
TournAr  
chMM0

liegement  
1242/01/01 ChMM019 7  
scriptorium: STilChatel? AbbLorr?  
charte: lettres réversales  
rante livres de terre // en la terre  
de Tyle Chastel // del duc // pour  
moi // et pour mes // hors // en tel  
maniere // que je et mes hors //  
après // moi , c' il qui iert syres //  
de // Tyle Chastel // tenra //

liegement del duc // et des ces hors  
// après lu // ces //  
quarante livres de terre  
davant dites // par en si  
cum ces leitres le //  
devisent //. Et si doé  
asennee // cele terre //; et  
quant je l' avrae asennee  
je doé doner mes

◆ 1° “ [Ouvrir ce passage dans DocLing](#) ”

13 5; chCOR032 6; ChMa132 4;  
140b, *ligement* BenTroieC;  
X, 26; GaceBruléD \*6,20; Gloss.;  
CoutVerdun<sup>1</sup>M Anm. S. 152 zu Vers  
M040 9; ChMM053 4; ChMM058  
; ChMM087 4; ChMM087 5;  
4214 3; GilebBernH S. 80 v. 4 s. Ag.  
ChMM279 10; ChMM285 3;  
ocJuBe017 8; chN 023 3;  
ynMotets I, p. 52, etc.;  
GodinM 8505, 11428, 16864;  
m.); ClefD 2510, *lygement*  
9241; BenDucM I, p. 486, v. 1172;  
AlexParA (ß 87) 1403 var. t. 5; IpH 7633; CoucyChansL XXX, 26; GaceBruléD \*6,20; Gloss.; GaceBruléH; PrêtreJeann/vG  
606; MerlinM; Bueve3S 14484(s. Anm.); DurmG gloss.; MoniotArrD v. gloss.; DolopL; GodinM 8505, 11428, 16864;  
GuillVinM II, 14; ChansArtB XVIII, 189; CoutVerdun<sup>1</sup>M Anm. S. 152 zu Vers 22,60; DocHMarneG 124, 6; TournAntW 1688;  
DocHainR 8; 2; BestAmOctT 10, 508; v. gloss.; HistFécl. 2249; JeuxPartL; GilebBernF; GilebBernH S. 80 v. 4 s. Ag. W;  
EtSLouisV II, 396; ClefD 2510; Acarth 1802; TombChartr1/2/3S; [ProstInv 519; DeschQ VI, 126 etc.]; (sigles à datations  
multiples:) RaynMotets I, p. 52, etc.; SchwanBehrens p. 140b; StimminMot, ContPerc1R T 12338; CorleyCont<sup>2</sup> 28291;  
Morlet 273; MéiMel 221; Runk p. 108; SchelerJPreis 30901 etc.; [sigle], Gdf IV 781a; Gdf s. *ligement*; AND 388a; AND 382a;  
DCCarp 253c; v. gloss.; FEW s. *linge*, adj.; DupinLaboulaye)

### III. 12. Contexte court d'une occurrence des DocLing et hyperliens au texte intégral

À partir d'ici, un deuxième hyperlien amène à l'édition électronique de l'acte où l'occurrence en question est marquée en gras :

**Accueil**

**Descriptif du projet**

**Les corpus textuels**

Les corpus textuels ...

Table des dates

[actualiser](#)

**Interrogations linguistiques**

[\[vers une autre interface\]](#)

**Contact**

**Choix d'édition**

☒ édition critique

☐ édition interprétative

☐ édition diplomatique

**Montrer la structure rhétorique**

☒ non

☐ oui

[actualiser](#)

**Téléchargement du document**

[fichier xml](#)

[fichier pdf](#)

Documents linguistiques galloromans  
Corpus : chartes de Meurthe-et-Moselle (ChMM)  
Responsable du corpus : Martin-D. Glessgen  
Édition de la charte : Martin-D. Glessgen / Michel Arnold (révision: D. Kihal)

**ChMMo62**

1248, septembre

TYPE DE DOCUMENT: charte: lettres réversales (cf 065)  
OBJET: Acte correspondant au précédent, établi au nom et sous le sceau de Mathieu [II] duc de Lorraine.

AUTEUR: Mathieu [II] duc de Lorraine  
SCAUX: AUTEUR  
BENEFICIAIRE: Mathieu [II] duc de Lorraine  
RÉDACTEUR: chancellerie de Mathieu [II] duc de Lorraine [pal]

SUPPORT: parchemin soigné sur double queue  
LIEU DE CONSERVATION: AD MM B 793, n° 4, pièce 4, layette Lunéville  
VERGÉ: Lettre d'ou duch pour Dongewin (XIV<sup>e</sup> s.)

1 Je, Maheus, dux de Lohetaine et marchis, 2 faiz connoissant à toz 3 que Ferris de Salmes et sires de Blamont, mes coisins, est devenus mes hom luges après mon 4 signor l'aveske de Mesz: envers toz homes- 4 et je li ai rendu le ban de Donjevin et quant <sup>(1)</sup> qui i apant et le ban de Leffenborne et quant <sup>(2)</sup> qui i-apant: et ce 5 tient il de moi ligement: 5 et 6 at acreu le fît devant nomé de Frisonviler et del molin qui <sup>(3)</sup> estoit ses alués; 6 Et est asavoir qu'après <sup>(4)</sup> son décès, 7 il uns de ses oirs de son propre cors masles: se plus en at d'un après le premerain masle, sera mes hom devant toz homes: et tanra les choses devant nomees de moi et des miens ensi com il les en tient:

7 Et se il n'avoit nul oir masle mais c'un- et il aust femele; la fe-mele le tanroit en-tel point com li 8 masles, ensi com il est deviz; 8 et se il n'avoit mais c'un oir de son propre cors, fust masles, fust 9 femele: adés seroit il mes hom luges après mon signor l'aveske de Mes; et tanroit ces choses devant dites ensi com il est deviz:

9 Et totes les fois qu'il avroit dous oirs à Blamont: adés en seroit li uns mes hom luges et à mes oirs: envers toz 10 homes ensi com il est deviz: et tanroit les choses devant dites ensi com il est deviz: 10 et se il oir ne voloient ce faire, li fies 11 devant només renvoit à moi et à mes oirs tot entierement, par son lous et par son creant;

### III. 13. Occurrence des DocLing dans son contexte d'origine

L'intégration des 4 300 lemmes est désormais achevée et, par cette voie, les données lexicales des DocLing sont enfin accessibles pour à la communauté scientifique sous une forme structurée.

## 3.4 L'intégration sémantique

L'intégration sémantique des lemmes est considérablement plus exigeante. Dans un premier temps, les occurrences des DocLing sont placées dans la partie des sens et syntagmes du DEAFpré sous une entrée à la fin des articles, appelée « Identificanda DocLing ». En attendant le traitement sémantique des formes par nous, les utilisateurs peuvent prendre connaissance des formes et, si nécessaire, en vérifier le sens en prenant en considération le contexte des documents. Les hyperliens fonctionnent aussi dans la partie sémantique.

Voici les « Identificanda » dans l'article *ligement* :

• **ligement** adv.

[Gdf IV 781a; Gdf s. *ligement*; AND 388a; AND 382a; DCCarp 253c; v. gloss.; FEW s. *linge*, adj.; DupinLaboulaye. – ContPerc1R T 12338; CorleyCont<sup>2</sup> 28291; Morlet 273; Mélmel 221; Runk p. 108; SchelerJPreis 30901 etc.; [sigle].]

(*lie/g*)*emant* chCor013 3, *liegement* ChMM019 7, *ligemant* DolopL; ChMM062 4; chCor013 5; chCor032 6; ChMa132 4; ChMM219 6; ChMa205 10, *ligemant*, adv. (sigles à datations multiples:) SchwanBehrens p. 140b, *ligement* BenTroieC; BenDucM I, p. 486, v. 1172; AlexParA (ß 87) 1403 var. t. 5; IpH 7633; CoucyChansL XXX, 26; GaceBruléD \*6,20; Gloss.; MerlinM; DurmG gloss.; MoniotArrD v. gloss.; GuillVinM II, 14; ChansArtB XVIII, 189; CoutVerdun<sup>1</sup>M Anm. S. 152 zu Vers 22,60; DocHMarneG 124, 6; DocHainR 8; 2; ChMe012 5; chMM014 4; chMe025 3; ChMM040 9; ChMM053 4; ChMM058 15; ChMM061 4; BestAmOctT 10, 508; v. gloss.; HistFécL 2249; JeuxPartL; ChMM078 3; ChMM087 4; ChMM087 5; chCor032 5; ChMM091 3; ChMM153 5; ChMM165 4; ChMM168 7; ChHM124 8; ChMM214 3; GilebBernH S. 80 v. 4 s. Ag. W; GilebBernF; ChMM223 10; ChMM238 3; ChMM238 4; ChMM238 4; ChMM246 4; ChMM279 10; ChMM285 3; chCor184 7; chMe237 4; EtSLouis V II, 396; chCor032a 12; chN 008 9; docJuBe017 8; docJuBe017 8; chN 023 3; TombChartr1/2/3S; [ProstInv 519; DeschQ VI, 126 etc.]; (sigles à datations multiples:) RaynMotets I, p. 52, etc.; StimmingMot, *ligement*, adv. BenDucF 13845, 29241; GaceBruléH; PrêtreJeann/vG 606; GodinM 8505, 11428, 16864; TournAntW 1688; AcartH 1802, *lingemant* ChMM233 3, *lingement* Bueve3S 14484(s. Anm.); ClefD 2510, *lygement* chMM023 19; chMM023 20; ChMM058 3)

◆ 1° “selon le lien féodal qui unit le vassal au seigneur” (BenTroieC; BenDucF 13845, 29241; BenDucM I, p. 486, v. 1172; AlexParA (ß 87) 1403 var. t. 5; IpH 7633; CoucyChansL XXX, 26; GaceBruléD \*6,20; Gloss.; GaceBruléH; PrêtreJeann/vG 606; MerlinM; Bueve3S 14484(s. Anm.); DurmG gloss.; MoniotArrD v. gloss.; DolopL; GodinM 8505, 11428, 16864; GuillVinM II, 14; ChansArtB XVIII, 189; CoutVerdun<sup>1</sup>M Anm. S. 152 zu Vers 22,60; DocHMarneG 124, 6; TournAntW 1688; DocHainR 8; 2; BestAmOctT 10, 508; v. gloss.; HistFécL 2249; JeuxPartL; GilebBernF; GilebBernH S. 80 v. 4 s. Ag. W; EtSLouis V II, 396; ClefD 2510; AcartH 1802; TombChartr1/2/3S; [ProstInv 519; DeschQ VI, 126 etc.]; (sigles à datations multiples:) RaynMotets I, p. 52, etc.; SchwanBehrens p. 140b; StimmingMot, ContPerc1R T 12338; CorleyCont<sup>2</sup> 28291; Morlet 273; Mélmel 221; Runk p. 108; SchelerJPreis 30901 etc.; [sigle], Gdf IV 781a; Gdf s. *ligement*; AND 388a; AND 382a; DCCarp 253c; v. gloss.; FEW s. *linge*, adj.; DupinLaboulaye)

◆ 2° “Identificanda DocLing” (ChMe012 5; chMM014 4; ChMM019 7; chMe025 3; chMM023 19; chMM023 20; ChMM040 9; ChMM053 4; ChMM061 4; ChMM062 4; ChMM058 3; ChMM058 15; chCor013 3; chCor013 5; ChMM078 3; ChMM087 4; ChMM087 5; ChMM091 3; chCor032 5; chCor032 6; ChMM153 5; ChMM165 4; ChMM168 7; ChHM124 8; ChMa132 4; ChMM214 3; ChMM219 6; ChMM223 10; ChMM233 3; ChMM238 3; ChMM238 4; ChMM238 4; ChMM246 4; ChMM279 10; ChMM285 3; chCor184 7; chMe237 4; ChMa205 10; chCor032a 12; chN 008 9; docJuBe017 8; docJuBe017 8; chN 023 3)

Cf. ANDEL *LIGEMENT*<sup>1</sup> □; DMF *LIGEMENT* □.

III. 14. Article en ligne du DEAFpré avec les occurrences des DocLing dans le répertoire formel et dans la partie sémantique

Pour mener à bien leur véritable intégration sémantique, il faut cerner le sens des différentes occurrences, dégager les syntagmes éventuels ou encore le cadre valentiel des verbes, souvent particulier dans les textes documentaires. Ensuite, les résultats sont intégrés dans la structure sémantico-syntagmatique de l'article du DEAFpré, ce qui suppose souvent des interventions dans cette structure, conçue par les rédacteurs du DEAF comme provisoire.



Voici les occurrences du lemme *laborage* avec leurs contextes dans le logiciel *Phoenix2*, que nous consultons de nouveau dans cette étape afin de dégager les sens et les syntagmes<sup>15</sup>.

Ill. 15. Variantes et occurrences d'un lemme donné dans le logiciel *Phoenix2* (cf. ill. 4)

Ensuite nous attribuons les attestations des DocLing à un sens préexistant dans le DEAFpré, ou, si nécessaire, nous ajoutons une nouvelle définition à l'aide des matériaux du FEW ou du DMF, en indiquant alors ces dictionnaires dans la définition :

Ill. 16. Structure sémantique arborescente dans le système de rédaction du DEAFpré

<sup>15</sup> Comme nous l'avons déjà mentionné, dans les modules de *Phoenix2*, les chartes concernées sont à tout moment accessibles par hyperlien dans les cas où le contexte restreint ne suffirait pas pour l'identification du sens.

Si nous avons été amenée à préciser l'arbre sémantique du DEAFpré, nous redistribuons dans la mesure du possible les fiches du DEAFpré, en fonction de leur nombre et de la précision des indications sémantiques fournies – comme on le sait, les fiches du DEAFpré sont des fiches de renvoi, comportant parfois des définitions, mais presque jamais des contextes.

L'intégration sémantique peut être considérée comme l'achèvement d'une description lexicologique simple des *Documents linguistiques*. Le résultat a alors toutes les qualités d'un glossaire traditionnel correctement réalisé, à la différence près que notre « glossaire » englobe une grande quantité de textes et non pas, comme d'habitude, un texte individuel ou un petit ensemble de textes. Dans ce sens, nous rejoignons ici la logique du Godefroy, qui est, lui aussi, plus un macro-glossaire qu'un dictionnaire philologique élaboré. Les autres avantages par rapport à un glossaire traditionnel résident dans les hyperliens, permettant l'accès direct aux contextes, et dans la présence parallèle des autres occurrences du DEAFpré. Celles-ci fournissent – au-delà des cautions émises par les rédacteurs – tous les renvois aux dictionnaires de référence ainsi que toutes les indications sémantiques des glossaires réunis dans le fichier du DEAF. Il est ainsi possible de se faire une idée de la fréquence d'une forme lexicale dans les textes médiévaux et éventuellement aussi de sa présence dans les différents genres textuels littéraires et scientifiques. Cette information est précieuse pour notre corpus parce qu'elle permet de distinguer des lexèmes qui appartiennent essentiellement à des textes documentaires de ceux qui connaissent une diffusion relativement équilibrée parmi les différents genres et ont donc une portée plus générale. Cela vaut notamment pour les mots traités dans le DEAFplus, qui représentent un segment significatif du vocabulaire dans lequel les occurrences des DocLing se placent dans un cadre bien établi et parfaitement structuré.

Nous avons entrepris les premiers sondages concernant l'intégration sémantique, qui nous occupera tout au long de l'année à venir, après les étapes préliminaires de l'importation et de l'intégration formelle. Une fois achevée, nous disposerons ainsi d'un macro-glossaire de toutes les séries des *Documents linguistiques* réunies dans l'édition électronique, comme nous l'avions envisagé au début. Par ailleurs, les 4 300 articles concernés du DEAFpré connaîtront par là une révision, au moins partielle, qui n'équivaut bien entendu pas le traitement des 9 000 articles du DEAFplus, mais qui représente néanmoins une plus-value réelle pour les articles provisoires en ligne.

### 3.5 La rédaction d'articles lexicologiques dans le cadre du DEAFpré

Une fois ce premier objectif atteint, nous avons souhaité continuer dans cette voie, en exploitant la dimension évolutive du DEAFpré. Grâce à son système rédactionnel performant, il permet des élaborations lexicographiques achevées, ce qui nous a amené à développer ultérieurement les entrées concernées par les *Documents linguistiques*.



Notre idée est d'élaborer sous les données des DocLing la forme d'articles lexicologique, en exploitant les informations disponibles dans la lexicographie de référence et dans les fiches du DEAFpré. Il ne s'agit pas de rédiger des articles du DEAFpré dans la logique du DEAFplus, mais de cerner plus précisément les dimensions étymologiques, sémantiques, syntagmatiques et diasystématiques des occurrences de notre corpus. Cette rédaction se concentrera sur les mots que l'on peut considérer comme caractéristiques des textes documentaires. Cela concerne donc des mots ayant trait aux domaines sémantiques de l'agriculture, de la gestion patrimoniale, du droit privé ou de l'Église ou encore des mots qui montrent une concentration reconnaissable dans les genres textuels documentaires. Notre choix porte ainsi parallèlement sur les paramètres onomasiologiques et de l'usage. Par ailleurs, nous souhaiterions mettre en relief les mots régionaux, qui sont en cours de traitement dans notre *Dictionnaire des régionalismes du français médiéval* (DRFM), au moins pour les régions de l'Est et du Sud-Est oïliques.

Voici un exemple d'un article rédigé<sup>16</sup>:

DocLingGloss – Article élaboré sur la base des matériaux des DocLing comportant une révision (partielle) du DEAFpré (resp. M. Glessgen).

frcomt. **LENÇOLEE** f.

rédaçtion: Marguerite Dallas (DocLing)

[FEW 5,367a lt. \*LÎNTEÖLUM.]

(lençolees docJuBe147 7)

◆ Synt.: *lençolee de foin* f. "quantité de foin permettant le rembourrage d'une housse de matelas (ici: utilisé comme mesure de surface)" (docJuBe147 7)

Le lexème est employé dans un rôle de terres provenant des environs de Porrentruy pour désigner une mesure de terre (= la surface nécessaire pour produire la quantité de foin utile au rembourrage d'une housse de matelas), parallèlement à une 'journée de terre' ou une 'faux de pré'.

Le dérivé en -ATA de lt. LÎNTEÖLEUM est présent dans les dialectes de l'est de la Galloromania entre le gaumais et la Provence ainsi qu'en Gascogne pyrénéenne, toutefois avec une rupture de continuité entre le gaumais et le franc-comtois et avec une présence particulièrement nette en franc-comtois, francoprovençal et nord-provençal. Les seules attestations anciennes connues par le FEW sont apr.occ. *lansolada* (*Hist. de Nîmes* II, 319 = [Lv]) et alang.or. *lensolada* (1436, Montagnac / Hérault, FEW 5, 367a).

L'attestation du Jura bernois antédote la documentation et confirme l'ancienneté de la diffusion dialectale moderne qui semble indiquer une formation protoromane régionale du type \*LÎNTEÖLATA.

Le terme appartient dans l'usage aux genres documentaires. [MD/MG]

### III. 17. Article d'un lexème des DocLing rédigé dans le cadre du DEAFpré

<sup>16</sup> L'article indique en haut à droite le rédacteur/la rédactrice initial(e) de l'article et à la fin du commentaire toutes les personnes qui sont intervenues au cours de la rédaction.

Un deuxième exemple, l'article *leude*, illustre la possibilité d'intégrer des hyperliens complémentaires comme, ici, à la documentation des attestations latinisées identifiées, relevées dans le corpus des *Chartae Galliae* (cf. Carles 2017, 24–25) :

• **leude** f.

(*leuda* CptRoyM 3154; CptRoyM, *leude* RV 1292 07 25 01 6; RV 1292 07 25 01 6, *leudes* RV 1292 07 25 01 6)

- ◆ “droit d’importation et de vente levé sur les marchandises au sud de la Galloromania” (CptRoyM 3154; CptRoyM)
- ◆ *menues leudes* “droit d’importation et de vente levée sur les menues denrées” (RV 1292 07 25 01 6)
- ◆ *leude du sel, de l’avainne, dou blé* (RV 1292 07 25 01 6; RV 1292 07 25 01 6)

Le lexème est employé dans une lettre patente de la Chancellerie royale sous Philippe IV (‘le Bel’) pour désigner un type de redevance levée par l’administration d’une ville sur l’importation et sur la vente au marché et aux foires du sel et des céréales. La redevance est typique du sud de la Galloromania médiévale, correspondant au *tonlieu* au nord. Elle fait partie intégrale ici des possessions appartenant au château de Montferrand dans la région Puy-de-Dôme; le lexème francisé est donc employé ici exclusivement pour désigner une réalité appartenant au domaine occitan. Les redevances sur ces produits mentionnés constituent évidemment les ‘grosses leudes’ par opposition aux *menues leudes* également citées dans le document, tout comme aux ‘leudes mages’<sup>(1)</sup>

Il s’agit d’un emprunt à l’aocc. *leuda* (cf. Rn 4, 61b “droite de péage, sorte de tribut” et Lv 875a “droit d’importation ou de vente sur les marchandises”).

Le terme appartient dans l’usage aux genres documentaires.

Cf. ChGalliae [LEUDE](#) .

<sup>(1)</sup> Cf. Aigrefeuille 1877, vol. 2, 675: *Elles [= les leudes] étoient divisées en leudes grosses et leudes menues ainsi nommées pour les distinguer de la leude mage, qui consiste dans les droits exigez pour le poids du roi.*

### III. 18. Deuxième exemple d’un article rédigé avec un hyperlien aux *Chartae Galliae*

Dans le cas où un mot a déjà été traité dans le DRFM, notre traitement se limitera à citer ce dernier<sup>17</sup> :

<sup>17</sup> Les sigles AB, HC, MG, GR et MR se réfèrent aux rédacteurs concernés du DRFM, Alessandra Bosson, Hélène Carles, Martin Glessgen, Gilles Roques et Marco Robecchi.

- champ. lorr. bourg. **leschiere** f.

[Gdf IV 761b; FEW V 372b; FEW V 372b. – SturmSumpf p. 70.]

(*lechiere* ChMa119 10; ChMa119 18, *Lechierre* ChV0101 5, *leschere* SimPouilleAB B, *leschiere* ChV0145 2; ChV0145 3, *Liechiere* ChV0099 3, *lischiere* ChMa119 13, *loichere* RegDijon<sup>1</sup>L)

- ◆ 1° "prairie marécageuse où poussent les laïches" (ChMa119 10; ChMa119 13; ChMa119 18; ChV0099 3; ChV0101 5; ChV0145 2; ChV0145 3; RegDijon<sup>1</sup>L, SturmSumpf p. 70, Gdf IV 761b; FEW V 372b; FEW V 372b)
- ◆ 2° "plante qui pousse au bord de l'eau, roseau" (SimPouilleAB B)

**DRFM s.v. *leschiere*** Le type lexical est connu dans l'aire oïlique sud-orientale, notamment en Champagne, Lorraine et Bourgogne entre 12e et 14e s. Il survit dans les parlers modernes de l'aire frpr. : Bresse, Savoie, Rhône et Dauphiné.

Dans la plupart des occurrences le substantif semble désigner un lieu où poussent les laïches (carex), c'est à dire une prairie marécageuse. Le sens "roseau" est connu par une seule attestation littéraire. Par ailleurs, le type lexical s'est aussi cristallisé comme toponyme avec et sans article. Le type sans article est présent dans les régions Champagne, Lorraine et Bourgogne (dp. 10e s.) ; en revanche, celui à article est plus tardif et est attesté essentiellement en domaine frpr. (dp. 13e s.) et en Bourgogne (v. Nègre 1729 et 1725 : La Lochère).

La majorité des occurrences est tirée de textes documentaires ; la seule attestation littéraire est tirée de Simon de Pouille, il s'agit de la leçon du ms lorrain BN fr. 368 et illustre, quant à elle, un sens différent.

Il s'agit d'un continuateur du protoroman \*LISCARIA, formé au moyen du suffixe -aria sur le substantif protoroman \*lisca "laïches" qui est, quant à lui, d'origine prélatine (mais cf. lat. méd. *lisca* dp. 9e s., Vitali). La présence de nombreuses formes toponymiques sans article nous permet de supposer une formation ancienne. Le sens 2° se dégage du premier par métonymie suite à une relation partie-tout.

La diffusion du régionalisme concerne l'aire oïlique sud-orientale : la Lorraine, la Champagne et ponctuellement la Bourgogne ; par ailleurs, les données toponymiques et dialectales nous permettent d'élargir sa présence à l'aire frpr. Notons que les continuateurs du protoroman \*LISCA montrent, quant à eux, une diffusion plus ample (toute la Galloromania ; cf. aussi it. *lisca*, kat. *llesca*, etc.) » [AB/HC]

### III. 19. Article s'appuyant sur le DRFM

Nous sommes encore en train d'élaborer avec l'équipe du DEAF un modèle pour des articles du DEAFplus où notre rédaction pourra se contenter d'un commentaire supplémentaire, mettant en relief la part des sources documentaires.

Nous envisageons de traiter sous cette forme environ 1 200 lexèmes, soit un quart des lemmes contenus dans les DocLing, sachant qu'environ 400 de ces mots ont été traités dans le DRFM et *ca* 100 supplémentaires dans le DEAFplus. Sur cette base, nous souhaiterions mener en synthèse une réflexion sur la constitution du vocabulaire caractéristique des textes documentaires médiévaux. Il s'agira là de la première étude holistique permettant de caractériser d'un point de vue lexical un genre textuel donné.

## 4 Les DocLing et le Godefroy

Qu'en est-il maintenant de notre idée initiale d'un « Nouveau Godefroy » ? La réponse est partagée. Dans un sens, notre apport reste faible : une fois notre réalisation achevée, elle ne se substituera nullement au remarquable dictionnaire de Frédéric

Godefroy. Notre nomenclature englobe moins de 10 % de la sienne et elle ne cible qu'un seul ensemble textuel défini.

Dans un autre sens, nous apportons néanmoins des éléments significatifs. Étant donné que dans les domaines littéraires et scientifiques, la matière du Godefroy a déjà connu un renouveau presque intégral, le vocabulaire des DocLing se place précisément dans le domaine où jusqu'ici, le Godefroy est resté la source de renseignement la plus fiable sinon unique. Il est vrai que même pour les textes documentaires, nous ne couvrons qu'une partie de la nomenclature, mais il s'agit alors d'une partie quantitativement importante.

Ajoutons que dans de nombreux cas, nos éditions et notre traitement apportent des corrections et des précisions aux occurrences réunies dans Godefroy. Voici quelques exemples :

(1) Ajout de nouveaux lemmes à la nomenclature :

- frcomt. *beveuge* s.m. 'lieu où on loge les bestiaux, étable' [DRFM, AB]
- apic. *boistiel* s.m. 'mesure pour les grains' [charte de Douai 473a]
- frcomt. *lençolee* s.f. [document du Jura Bernois, cf. *supra*]

(2) Antédation de premières attestations :

- Gdf 3, 64a *emplastre* s.m. 'espace de terrain considéré comme propre à y construire, emplacement, terrain bâti': 1309 → 1254 [= charte de la Haute-Marne 41 ; DRFM, AB]
- Gdf 4, 335c [= DEAF G 1213] *grangete* s.f. 'petite grange': 1374 → 1270 [charte de Douai 441]
- Gdf 10, 494c *rebouchier* v.tr. 'construction, réparation': 1412 → 1268 [charte de la Haute Marne 234]

(3) Correction d'erreurs de transcription (plus fréquentes dans les contextes d'un lemme que dans les lemmes eux-mêmes) :

- Gdf 1, 21c, s. v. *abesoignier* (l'exemple correspond à notre charte de Meurthe-et-Moselle 103 [1254]), *aiwagier* → *anwagier* [l'attestation est alors à ajouter à Gdf 9, 468b/c *enguagier*; cf. FEW 17, 444b, \**waddi*]

L'apport de la nomenclature des DocLing ressort très nettement de son traitement par le DRFM qui met en relief de très nombreux ajouts, corrections et précisions géo-chronologiques et – surtout – sémantiques par rapport au Godefroy ; quelques exemples :

- Gdf 1, 596a afrb. *bastube*, -a s.f. 'auberge où l'on couche ?' → DRFM frcomt. JuBe *badestoube* s.f. 'étuve, établissement de bains chauds' (emprunt à l'alémanique) [AB]<sup>18</sup>.
- Gdf 2, 185c *colonge* s.f. 'fonds possédé par le colon' (champ. 1300, frcomt. 1343, neuch. 1395) → DRFM champ., bourg., frcomt. JuBe *colonge* s.f. 'tenure concédée à un colon dépendant d'un même seigneur et ayant son organisation propre' (t. de droit féodal), 1264- [AB]

<sup>18</sup> L'article du DRFM permet également la correction des définitions pour les dérivés répertoriés par Gdf : *bastubeor* s.m. 'hôtelier ?' ; *bastuber* v. 'passer la nuit dans une auberge ?'.

- Gdf 4, 134b s.v. *frefel* (var. *frefeil*, *ferfel*, *fravall*) s.m. ‘trouble, tumulte, agitation’ (pic., frcomt., Frib.) et *fravall* ‘vol de bois’ (SR Frib) → DRFM pic., frcomt. JuBe *frefel* s.m. ‘état d’agitation et de tumulte ; discussion entre deux parties ; sorte de délit’ (et *être en f.* ‘être en état d’agitation, affairé’) (emprunts parallèles à l’alémanique et au flamand) [MR]
- Gdf 4, 637c s.v. *jarron* s.m. ‘branche d’arbre’ (bourg.) → DRFM bourg., lorr. *jarron* s.m. ‘grosse branche d’arbre pouvant être utilisée pour faire du feu ou comme gourdin, rondin’ [MR]
- Gdf 7, 624a *taienos* m. ‘petit neveu’ (lorr.) → DRFM lorr. sept. *taienot* s.m. ‘fils d’un de ses propres fils, petit-fils’ [MR]

Les données des DocLing permettent également de réorganiser la structure des entrées retenue par Gdf ; par exemple :

- Gdf 1, 494b s.v. *aubue* s.f. [sans définition], (bourg 1294) et Gdf 4, 458c s.v. *herbue* s.f. ‘terrain rempli d’herbes’ (champ. *arbue* 1264 [= charte de la Haute-Marne 184]) sont à réunir ; cf. DRFM lorr., champ., bourg., frcomt. *albue* s.f. ‘terre glaise de couleur grise ou rougeâtre, argileuse’ [AB/HC]
- Gdf 4, 323b *gouterot* s.m. ‘gouttière’ (bourg.), adj. ‘(mur) auquel une gouttière est appliquée’ (champ.) ; 4, 299a *gocterot* s.m. [sans définition] (bourg.) et 4, 324a *gouteret* s.m. ‘gouttière’ (lorr.) sont également à réunir ; cf. DRFM champ., lorr. occ., bourg. *goterot* adj./s.m. ‘petit toit établi pour éviter l’écoulement d’eaux pluviales sur la façade, avant-toit’ et (par métaphore) ‘partie antérieure d’un parent d’autel ou d’un dais, frange’ [AB/MG]
- Gdf 5, 57b s.v. *maceclier* s.m. ‘boucher, charcutier’ comporte correctement des occurrences qui sont à rattacher à *maçacrier* (< \*MATTEUCCULARE), mais aussi d’autres qui appartiennent à afr. *maselier* (< MACELLARIUS) traité Gdf 5, 194a ; les occs. flandr. 1284/85, frcomt. 15<sup>e</sup> s., mfr. 1309, 1454 s.v. *maceclier* sont donc à déplacer s.v. *maselier* ; cf. DRFM s.v. pic., champ., bourg., frcomt. *maçacrier* [AB/MG]

Grâce à l’apport complémentaire du DRFM, les avancées par rapport au Gdf et à la lexicographie du français médiéval sont particulièrement nettes dans le domaine des régionalismes. Lors de la rédaction des articles dans le cadre du DEAFpré (cf. *supra* 3.5), cet aspect pourra surtout être exploité pour les régionalismes picards, très présents dans le volumineux corpus de Douai, et laissé de côté par le DRFM, étant donné sa concentration sur le domaine d’oïl sud-oriental.

L’importance de la nouvelle documentation des DocLing réside ainsi moins dans un élargissement de la nomenclature que dans un traitement adéquat d’un segment de vocabulaire jusqu’ici négligé. Si la formulation nous est permise : on connaissait les mots en question, mais on ne savait pas ce qu’ils voulaient dire.

## 5 Conclusions et perspectives

Finissons par trois réflexions de type méthodologique. Tout d’abord il nous semble important de souligner que notre réalisation est évolutive. Nous pourrions ainsi traiter dans les années à venir la matière lexicale provenant du territoire francoprovençal, déjà intégrée dans les DocLing. Celle-ci est également riche en régionalismes et com-

porte en ligne générale une nomenclature plus large que les documents du territoire d'oïl parce qu'il s'agit plus souvent d'inventaires et de listes que de chartes discursives.

D'autres corpus documentaires, comme celui du Doubs, sont encore en chantier et il est également possible d'envisager l'intégration de textes appartenant à d'autres genres. Le passé nous a appris à être prudent quant aux perspectives, mais celles-ci sont parfaitement réalistes et se placent, justement, dans la logique d'une lexicographie évolutive. Les étapes parcourues jusqu'ici avec la constitution d'un système informatique complexe et désormais stable permettent de se concentrer à l'avenir sur le travail philologique et lexicologique plus que sur l'élaboration infrastructurelle.

La deuxième réflexion est plus générale et concerne la nature de la lexicographie sur corpus. Le cas des DocLing prouve que, si l'on reste dans le cadre d'un genre textuel et d'une époque donnée, la variance lexicale n'est pas nécessairement immense. Il faut un nombre très important de ressources pour élargir la nomenclature. Dans le cas des chartes, on peut considérer que chaque nouvelle charte ajoutée apporte en moyenne un mot nouveau à l'inventaire constitué. Ce constat rejoint des observations que nous avons pu faire sur les corpus des lettres de la première Guerre mondiale exploités par Pierre Rézeau dans les *Mots des Poilus* (2018) ou encore d'autres qui ressortent de l'analyse du de l'ARTEM par Hélène Carles (2017). P. Rézeau a relevé dans quelque 100 000 lettres environ 15 000 occurrences concernant 4 000 lexèmes mal décrits par la lexicographie : même si sa récolte apporte un regard très nouveau sur cet épisode crucial dans l'histoire, il aurait été parfaitement démesuré de concevoir ce dictionnaire sur la base d'un corpus informatique. Le constat est identique pour le *Trésor galloroman des origines* (TGO) : H. Carles a traité 400 lexèmes vernaculaires relevés dans les 4 500 actes latins originaux édités ; elle aurait pu ajouter un certain nombre d'autres entrées, mais sa nomenclature n'aurait pas pu dépasser 1 000 entrées pertinentes. Si l'on considère que l'ARTEM a mis – avec un volume de quatre postes CNRS – 40 ans pour établir sa base textuelle, il devient évident que la constitution de cette base n'aurait pas pu se justifier par le seul objectif de traiter 400 à 1 000 lexèmes vernaculaires, aussi intéressants soient-ils pour mieux cerner les processus de fragmentation et de transformation lexicale dans la Galloromania. Entendons-nous : de grandes entreprises nationales, comme le TLF pour le français moderne, basé sur Frantext, ou le TLIO pour l'ancien italien, basé sur l'OVI, ont montré le potentiel remarquable d'une lexicographie de référence sur corpus – mais chaque fois avec un investissement en termes de personnel qui se calcule en centaines et parfois en milliers d'années de travail. Les deux exemples des *Mots des Poilus* et du TGO illustrent en revanche que la méthode traditionnelle consistant à relever des mots dans des sources diverses sans forcément en fournir l'édition et sans les constituer en corpus peut donner des résultats très importants en termes de nomenclature.

Nous ne regrettons pas notre approche maximaliste, qui consiste à éditer un nombre important de textes nouvellement pour les exploiter ensuite dans une logique lexicographique. Au delà du fait que ce corpus permet également des interrogations

scriptologiques, morpho-syntaxiques, onomastique et historiques, il comporte d'un point de vue purement lexicologique une plus-value importante par ses contextualisations et par la sécurité philologique qu'il peut garantir. Dans le cadre d'une lexicographie pleinement informatisée, il apporte également un enrichissement à la documentation préexistante, en précisant l'usage, la fréquence et la distribution diasystématique des lexèmes en question.

Mais il nous semble important de combiner différentes approches en lexicographie diachronique. Dans cet esprit, nous avons intégré avec H. Carles dans les DocLing le corpus des documents provenant du territoire francoprovençal comportant des éditions de qualité, mais intégralement menés à bien par d'autres éditeurs. Et dans ce même esprit, il est imaginable d'enrichir le DEAFpré avec d'autres ressources textuelles informatisées. L'exemple des hyperliens aux *Chartae Galliae* montre qu'il est facile d'intégrer dans les articles des renvois à d'autres bases ou éditions de textes de référence.

Enfin : l'idée d'intégrer les mots des DocLing dans le DEAFpré, qui a coûté des énergies considérables autant à Sabine Tittel et ses coéquipiers qu'à nous-même, nous semble malgré ses lourdeurs la voie à suivre à l'avenir<sup>19</sup>. L'intégration des DocLing conduit à une amélioration sensible du DEAFpré et place en même temps sa documentation dans un cadre de référence qui la rend plus significative. Par ailleurs, le DEAFpré s'ouvre à la communauté scientifique et fournit ainsi une aide précieuse à l'orientation dans la trajectoire des mots. Dans ce sens, il nous semble indispensable d'instaurer des liens étroits entre la philologie éditoriale, la linguistique de corpus et la lexicographie électronique et de réduire l'éclatement des réalisations individuelles au profit d'un renforcement des ouvrages de référence et des connaissances partagées.

---

**19** Sur ce point, nous ne pouvons que nous rallier aux conclusions de S. Tittel (2018, 464 s.) : « To the best of our knowledge, this is the only example of a retrospective and successful integration of two voluminous and longstanding projects of a historical (Romance) language both from a graphical and a semantic point of view. (...) This [= the characteristics of our project] clearly contradicts Asmussen 2013 : 1087 : « Combining existing dictionaries with existing corpora will inevitably yield products of second quality ». Also, we conclude that the integration of DocLing and DEAF is a promising pilot project for the integration of other corpus linguistic data into a dictionary. At the same time, it emphasizes the role of the DEAF as a standard reference that can also be used for other single scholarly editions of Old French texts that are digitally published ».

## 6 Références bibliographiques

Les sigles non reproduits sont ceux de la DEAFBibl.

### 6.1 Monographies et articles

- Asmussen, J. (2013): Combined products: Dictionary and Corpus. In: Gouws, Rufus H./Heid, Ulrich/Schweickard, Wolfgang/Wiegand, Herbert Ernst (edd.), *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography* (HSK 5.4). Berlin / Boston: De Gruyter, 1081–1090.
- Büchi, Eva (1996): *Les structures du Französisches Etymologisches Wörterbuch*, Tübingen: Niemeyer.
- Carles, Hélène (2017): *Trésor galloroman des origines. Les trajectoires étymologiques et géolinguistiques du lexique galloroman en contexte latin (ca 800–1120)*, Strasbourg: ÉLiPhi.
- Carles, Hélène/Glessgen, Martin (2019): L'élaboration scripturale du francoprovençal au Moyen Âge. In: *Zeitschrift für romanische Philologie* 135, 68–157.
- Carles, Hélène/Dallas, Marguerite/Glessgen, Martin/Thibault, André (2019): *Französisches Etymologisches Wörterbuch. Guide d'utilisation*, Strasbourg: ÉLiPhi.
- Chauveau, Jean-Paul (2003): L'utilisation du *Dictionnaire* de F. Godefroy dans le FEW. In: Frédéric Duval (ed.), *Frédéric Godefroy. Actes du X<sup>e</sup> colloque international sur le moyen français*, Paris: Ecole des Chartes, 323–344.
- Glessgen, Martin (1989): « *Lo Thesaur del hospital de Sant Sperit* ». *Edition eines Marseiller Urkundeninventars (1399–1511)*, Tübingen: Niemeyer.
- Glessgen, Martin (1996): *Die Falkenheilkunde des « Moamin » im Spiegel ihrer volgarizzamenti. Studien zur Romania Arabica*. 2 vols, Tübingen: Niemeyer.
- Glessgen, Martin (2001): Das altfranzösische Geschäftsschrifttum in Oberlothringen : Quellenlage und Deutungsansätze, In: Gärtner, Kurt et al. (edd.), *Skripta, Schreiblandschaften und Standardisierungstendenzen*. Beiträge zum Kolloquium vom 16. bis 18. September 1998 in Trier, Trier: THF, 257–294.
- Glessgen, Martin (2003): L'élaboration philologique et l'étude lexicologique des *Plus anciens documents linguistiques de la France* à l'aide de l'informatique. In: Duval, Frédéric (ed.), *Frédéric Godefroy. Actes du X<sup>e</sup> colloque international sur le moyen français*, Paris: Ecole des Chartes, 371–386.
- Glessgen, Martin (2015): L'écrit documentaire et le projet des *Plus anciens documents linguistiques de la France*, In: Trotter, David (ed.): *Manuel de la philologie de l'édition* (Manuals of Romance Linguistics/MRL, 4), Berlin/Boston: De Gruyter, 267–295.
- Glessgen, Martin (2018): Les *Documents linguistiques* de la France. Histoire, présent et perspectives d'un projet centenaire. In: *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres* (CRAI) 161 [Juillet à Octobre 2017], 1239–1270.
- Glessgen, Martin/Alletsgruber, Julia (2010): Structure de la base de données lexicologique des *Plus anciens documents*. In: *Bibliothèque de l'École des Chartes* 168 [2011], 83–94.
- Glessgen, Martin/Kopp, Matthias (2005): Linguistic annotation of texts in non-standardized languages: the program procedures of the tool PHOENIX. In: Pusch, Claus / Kabatek, Johannes / Raible, Wolfgang (edd.), *Romance Corpus Linguistics II: Corpora and Diachronic Linguistics*, Tübingen: Narr, 147–154.
- Glessgen, Martin/Tittel, Sabine (2018): Le *Dictionnaire d'ancien gascon* électronique (DAGél). In: Antonelli, Roberto/Glessgen, Martin/Videsott, Paul (edd.): *Atti del XXVIII<sup>o</sup> Congresso Inter-*



- nazionale di Linguistica e Filologia Romanza* (Roma, 2016), Strasbourg: SLR/ÉLiPhi, vol. 1, 805–818.
- Möhren, Frankwalt (2015): L'art du glossaire d'édition. In: Trotter, David (ed.), *Manuel de la philologie de l'édition*, Berlin: De Gruyter, 397–437.
- Rézeau, Pierre (2018): *Les mots des Poilus*, Strasbourg: SLR/ÉLiPhi.
- Tittel, Sabine (2010): Le « DEAF électronique » – un avenir pour la lexicographie. In: *Revue de Linguistique Romane*, 74, 301–311.
- Tittel, Sabine (2018): Historical Corpus and Historical Dictionary: Merging two Ongoing Projects of Old French by Integrating their Editing Systems. In: Čibej, Jaka et al. (edd.), *Proceedings of the XVIIIth EURALEX International Congress: Lexicography in Global Contexts*, 17–21 July, Ljubljana, Ljubljana: University Press, 453–465. [DOI: 10.13140/RG.2.2.22835.12328]
- Tittel, Sabine (2019): *Integration von historischer lexikalischer Semantik und Ontologien in den Digital Humanities. Mit Materialien des Altfranzösischen, Mittelfranzösischen und Altgaskognischen als Linked Open Data ins Semantic Web*, Heidelberg, ms.

## 6.2 Dictionnaires et bases de données

Tous les sites ont été visités le 10 août 2019.

- BDSud = *Biblioteca Digitale del Sud*, Dipartimento di Studi Umanistici dell'Università del Salento/ Dipartimento Giuridico dell'Università degli Studi del Molise, 2009–2018. Publication électronique <<https://www.bdsud.it/>>.
- BFM = *Base du français médiéval 2016: une base de référence de sources médiévales ouverte et libre au service de la communauté scientifique*, dirigée par Guillot-Barbance, Céline/Heiden, Serge/Lavrentiev, Alexei, 2017. Publication électronique <<http://bfm.ens-lyon.fr>>.
- CORDE = *Corpus diacrónico del español*, Real Academia Española. Publication électronique <http://www.rae.es>.
- DAGél = *Dictionnaire d'ancien gascon électronique (DAGél)*, dirigé par Glessgen, Martin. Publication électronique: site interne (2016) <<https://dag.adw.uni-heidelberg.de>>.
- DEAF = *Dictionnaire étymologique de l'ancien français – DEAF*, fondé par Baldinger, Kurt, continué par Möhren, Frankwalt, puis Städtler, Thomas, Québec/Tübingen/Berlin: Presses de L'Université Laval/Niemeyer/De Gruyter, 1971– [= DEAFplus].
- DEAFBibl = Möhren, Frankwalt, *Dictionnaire étymologique de l'ancien français. Complément bibliographique*, Berlin/Boston, De Gruyter, 2016.
- DEAFpré = *Dictionnaire étymologique de l'ancien français*. Matériaux lemmatisés en ligne, préparés par l'équipe du DEAF. Publication électronique <<http://deaf-server.adw.uni-heidelberg.de>>.
- DocLing = *Documents linguistiques galloromans (DocLing). Édition électronique*, dirigée par Glessgen, Martin, en partenariat avec Carles, Hélène/Duval, Frédéric/Videsott, Paul. Troisième édition revue et fortement élargie (2016). Publication électronique [www.rose.uzh.ch/docling](http://www.rose.uzh.ch/docling).
- DRFM = *Dictionnaire des régionalismes du français médiéval*, dirigé par Carles, Hélène/Glessgen, Martin, (en prép.).
- GdfBibl = Ringenbach, Jean-Loup, *Bibliographie des sources du Godefroy*. Publication électronique [www.atilf.fr/BbgGdf](http://www.atilf.fr/BbgGdf).
- OVI = *Opera del Vocabolario Italiano*, Firenze: Consiglio nazionale delle ricerche/Accademia della Crusca, 1965–. Publication électronique <<http://www.ovi.cnr.it/index.php/it/il-corpus-testuale>>.

Steinfeld 2005– = Steinfeld, Nadine et al., *Base des mots fantômes (base dédiée aux « mots fantômes »: pseudo-lexèmes disposant à tort d'un statut lexicographique)*, Nancy, ATILF-CNRS. Publication électronique [www.atilf.fr/MotsFantomes](http://www.atilf.fr/MotsFantomes).

TLFÉtym = Steinfeld, Nadine, *Programme de recherche TLF-Étym (révision sélective des notices étymologiques du « Trésor de la langue française informatisé »*, Nancy, ATILF-CNRS et Université de Lorraine, 2005–. Publication électronique [www.atilf.fr](http://www.atilf.fr), onglet « Les projets ».